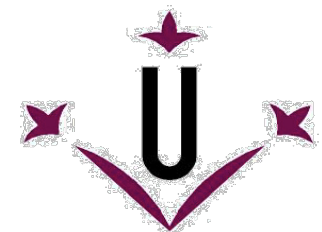


Enginyeria del Software III

XML XHTML

Roberto García
Despatx EPS 3.15
rgarcia@diei.udl.cat



INTRODUCCIÓ

○ HTML 4.0

- Llenguatge estàndard de publicació del WWW.
- Conforme amb l'estàndard ISO 8879.
- El seu èxit: basar-se en un conjunt molt reduït d'etiquetes estructurals de SGML.
- Molt estès (popular), però:
 - NO EXTENSIBLE → Necessitat constant d'inventar nous elements ...
 - No té perquè estar ben format ...
 - Accés als documents HTML des d'una enorme heterogeneïtat de plataformes ...
 - ... que encara causen innumerables problemes!!



INTRODUCCIÓ

○ XHTML 1.0

- És una reformulació de HTML com aplicació XML per a ser utilitzat com llenguatge de continguts conforme XML i que funcioni amb agents HTML.
 - <http://www.w3.org/TR/xhtml1>
 - <http://www.sidar.org/recur/desdi/traduc/es/xhtml11.htm>



XHTML

CARACTERÍSTIQUES

- Documents conformes amb XML
 - Fàcilment visualitzats, editats i validats amb eines XML estàndard.
- Poden utilitzar aplicacions (scripts, applets,...) basades en el **Document Object Model** (DOM) de HTML i de XML.
 - Objectiu DOM: definir una **interfície programable** per a HTML i XML.
- A l'igual que HTML, és portable.
- Però com també és XML, és **ampliable**.
 - Els desenvolupadors tenen major flexibilitat.
 - Poden crear les seves pròpies etiquetes i/o atributs.



XHTML

CARACTERÍSTIQUES

- A diferència de HTML, XHTML (per ser XML) es pot **analitzar sintàcticament**.
- Per ser XML, els elements són **'case-sensitive'**
 - <DISC> i <disc> són elements diferents.
- Tots els **elements** tenen que estar **correctament tancats** i correctament **niats**.
- Alguns **elements** (html, body, header, etc.) són **obligatoris**.
- OBLIGATORI: sempre s'ha d'incloure una **DTD**.



XHTML

DIFERÈNCIES AMB HTML

- De forma genèrica:
 - Compleix la “**regla document**” (RECORDAR: transparència següent).
 - Respecta totes les restriccions de la bona formació que es donen en l'especificació.
 - Cada una de les entitats analitzades està ben formada.
- Els **noms** dels **elements** i dels **atributs** han d'estar en **minúscules**
 - Perquè les DTD de XHTML estan en minúscules (<http://www.w3.org/TR/xhtml1/#dtds>)
<BODY BgColor="#000000"> és correcte en HTML ...
...però en XHTML ha de ser <body bgcolor="#000000">
- Els **valors** dels **atributs** van entre **cometes**
<table border=2> és correcte en HTML ...
...però en XHTML ha de ser <table border="2">



RECORDEM: REGLA “DOCUMENT”, DOCUMENTS BEN FORMATS XML

- Conté **un o més elements**.
- Hi ha exactament **un element**, anomenat **arrel**.
- Per a la resta d'elements,
 - Si l'**etiqueta** de **començament** està en el contingut d'algun altre element, l'etiqueta de **fi** està en el contingut del **mateix element**.
 - Es a dir, els elements delimitats per etiquetes de principi i final es nien mútuament de forma adequada.



XHTML

DIFERÈNCIES AMB HTML

- Tots els elements NO BUIITS han d'estar entre les seves etiquetes de principi i de fi.
 - HTML: l'obertura de moltes etiquetes tanca de forma automàtica l'anterior.
- I els buits han de portar l'acabament corresponent.
 - `<hr/>` o `
` (`<hr></hr>` o `
</br>`)
 - Recordatori: els elements buits en XML es declaren en la DTD amb la paraula EMPTY.
- Els niuaments han de seguir l'ordre d'obertura
 - `<p>text en negreta <i>i cursiva </i></p>`
és correcte en HTML ...
 - ...però en XHTML ha de ser
`<p>text en negreta <i>i cursiva </i></p>`



XHTML

DIFERÈNCIES AMB HTML

- XML, i per extensió XHTML, no suporta la minimització d'atributs

HTML

```
<OPTION VALUE="valor" SELECTED>  
<INPUT TYPE="tipus" CHECKED>  
<DL COMPACT>
```

XHTML

```
<option value="valor" selected="selected">  
<input type="tipus" checked="checked">  
<dl compact="compact">
```



XHTML DIFERÈNCIES CON HTML

- Definició dels elements style i script
 - Es declaren com elements CDATA

HTML

```
<SCRIPT LANGUAGE="JavaScript">  
document.write("hola, com estàs?");  
</SCRIPT>
```

XHTML

```
<script type="text/javascript"  
language="JavaScript">  
<![CDATA [  
document.write("hola, com estàs?");]]>  
</script>
```

- Encara que el més apropiat és incloure el codi en arxius externs

```
<script type="text/javascript"  
language="JavaScript" src="codi.js">  
</script>
```



XHTML

NOVETATS RESPECTE HTML

- Els documents XHTML han d'incloure una declaració de “tipus de document”.
- Ha de ser vàlid per una d'aquestes 3 DTDs públiques
 - Una DTD estricta: **strict** DTD
 - Una de moderada: **transitional** DTD
 - O una que permet dividir pàgines en marcs: **frameset** DTD
- També hi han DTDs per entorns específics:
 - Mòvils: XHTML Basic 1.1
<http://www.w3.org/MarkUp/DTD/xhtml-basic11.dtd>
 - Impressores senzilles: XHTML Print 1.0
<http://www.w3.org/TR/xhtml-print/DTD/xhtml-print10.dtd>



XHTML

NOVETATS RESPECTE HTML

o Strict DTD

- Declaració

```
<!DOCTYPE html  
PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

- Imposa restriccions molt dures
 - Per exemple no permet col·locar text que no estigui en elements de bloc como <p> o <div>.
- S'utilitza quan utilitzem CSS per “formatar”
- Aquesta DTD és la recomanada.



XHTML

NOVETATS RESPECTE HTML

o Transitional DTD

- Declaració

```
<!DOCTYPE html  
PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

- S'utilitza quan volem utilitzar informació d'estil directament al XHTML
- Menys restrictiu
 - Afegeix elements de presentació HTML a la DTD strict



XHTML

NOVETATS RESPECTE HTML

o Frameset DTD

- Declaració

```
<!DOCTYPE html
```

```
PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
```

```
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
```

- S'utilitza quan els documents incorporen marcs (frames)
- Advertència: per raons d'accessibilitat (entre d'altres) es desaconsella l'ús dels frames.



XHTML

NOVETATS RESPECTE HTML

- En XHTML l'element arrel sempre és `<html>`
 - En el qual es declara el namespace
<http://www.w3.org/1999/xhtml>
- Els atributs `lang` i `xml:lang`
 - Tot i que són opcionals es recomanable utilitzar-los quan es vulgui especificar l'idioma d'un element
 - El valor de l'atribut `xml:lang` té preferència.
 - Per tant seran vàlids:
`<html xmlns="http://www.w3.org/1999/xhtml">`
`<html xmlns="http://www.w3.org/1999/xhtml" lang="ca">`
`<html xmlns="http://www.w3.org/1999/xhtml" lang="ca" xml:lang="ca">`
 - La declaració de DTD va abans de l'element arrel.



XHTML. NOVETATS RESPECTE HTML

○ Exemple

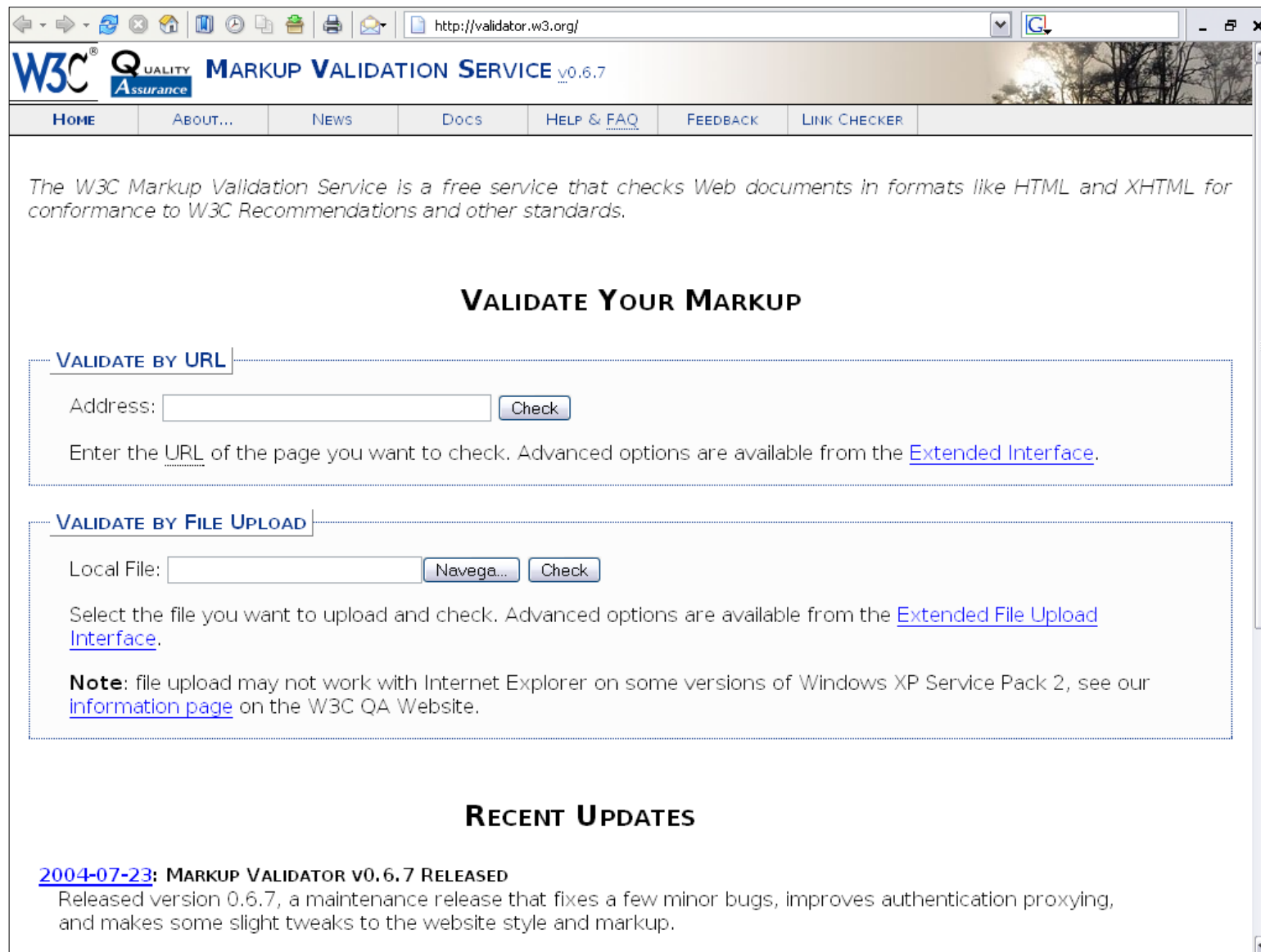
```
<?xml version="1.0" encoding="UTF-8"?>

<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">

<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="ca">
  <head>
    <title>Biblioteca Virtual</title>
  </head>
  <body>
    <p>Traslladada a <a href="http://vlib.org/">vlib.org</a>.</p>
  </body>
</html>
```



XHTML VALIDACIÓ (http://validator.w3.org)



The screenshot shows the W3C Markup Validation Service website. The browser address bar displays "http://validator.w3.org/". The page header includes the W3C logo and "QUALITY Assurance MARKUP VALIDATION SERVICE v0.6.7". A navigation menu contains links for HOME, ABOUT..., NEWS, Docs, HELP & FAQ, FEEDBACK, and LINK CHECKER. The main content area features a heading "VALIDATE YOUR MARKUP" and two validation methods: "VALIDATE BY URL" and "VALIDATE BY FILE UPLOAD". The "VALIDATE BY URL" section has an "Address:" input field and a "Check" button, with a note that advanced options are available from the "Extended Interface". The "VALIDATE BY FILE UPLOAD" section has a "Local File:" input field, a "Navega..." button, and a "Check" button, with a note that advanced options are available from the "Extended File Upload Interface". A "Note" section explains that file upload may not work with Internet Explorer on some versions of Windows XP Service Pack 2, with a link to an "information page" on the W3C QA Website. At the bottom, a "RECENT UPDATES" section lists the release of "MARKUP VALIDATOR v0.6.7" on 2004-07-23, describing it as a maintenance release that fixes minor bugs, improves authentication proxying, and makes slight tweaks to the website style and markup.



XHTML

CONVERTIR HTML A XHTML MANUALMENT

- Dos passos bàsics:
 - Assegurar-nos de que el document està ben format
 - Comprovar que es valida respecte a alguna de les 3 DTDs de XHTML
- Es poden dur a terme si se segueix la següent llista de comprovació (derivada de les diferències HTML - XHTML):
 1. Agregar una **declaració del tipus de document** que indica la DTD estàndard apropiada
 2. Declarar **l'espai de noms XHTML en l'element arrel**
 3. Assegurar-se que **tots els elements i atributs** estan definits **a la DTD**
 4. Posar tots els noms d'elements i atributs en **minúscules**
 5. Fer coincidir cada **etiqueta d'obertura** amb la corresponent **etiqueta de tancament**
 6. Acabar les **etiquetes buides** amb “/>”
 7. Incloure els **valors dels atributs** entre **cometes**
 8. Assignar **sempre valors als atributs**, encara que siguin cometes buides
 9. Definir els **valors dels atributs requerits** (obligatoris)



XHTML

CONVERTIR HTML A XHTML MANUALMENT

ACTIVITAT 1.6

- Convertir el document “html_INICIAL.htm”, què és HTML, a un document XHTML.

```
<html>
<script language="Javascript">
  function Hola(){alert ("Hola");}
</script>
<body>
  <h1>Pàgina HTML que convertirem a XHTML</h1>
  <p>Això és un paràgraf amb un text en <b>negreta <i> i</B> cursiva </i>.
  <p>Això és un paràgraf en el què hi col·loco un salt de línia. <br>
  I ara continuem aquí.</p>
  <P align=center>Això és text centrat</P>
  <ul>
    <li>Això és una llista amb un enllaç a una funció de
      <a href="javascript:Hola()">javascript</a></li>
    <li>Això és una altra llista amb una imatge 
    <li><font color=red size=4>Aquest és un altre element de la llista amb un altre
      color i mida de font major.</font>
  </ul>
  <p>I per acabar un formulari</p>
  <form>
    <input type="checkbox" checked>Opció 1
    <input type="checkbox">Opció 2
  </form>
</body>
</html>
```

XML. XHTML



XHTML

CONVERTIR HTML A XHTML AUTOMÀTICAMENT

ACTIVITAT 1.6

- Per convertir documents HTML a XHTML de forma automàtica podem utilitzar eines como Tidy
 - <http://www.w3.org/People/Raggett/tidy>
 - <http://tidy.sourceforge.net>
 - Eina en línia: <http://infohound.net/tidy>
- O el conversor en línia de la Universidad Carlos III de Madrid:
 - Eina en línia: <http://www.it.uc3m.es/jaf/html2xhtml/>
- **Activitat:** fer servir qualsevol de les eines en línia per la conversió manual anterior i comparar els resultats manual i automàtic.

XML. XHTML

